



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/044,782	01/11/2002	Andreas Arning	DE920000057US1	8377
25259	7590	02/01/2005	EXAMINER	
IBM CORPORATION			LY, ANH	
3039 CORNWALLIS RD.				
DEPT. T81 / B503, PO BOX 12195			ART UNIT	PAPER NUMBER
REASEARCH TRIANGLE PARK, NC 27709			2162	

DATE MAILED: 02/01/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)
	10/044,782	ARNING ET AL.
	Examiner	Art Unit
	Anh Ly	2162

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) Responsive to communication(s) filed on 30 August 2004.

2a) This action is **FINAL**. 2b) This action is non-final.

3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) Claim(s) 1-13 is/are pending in the application.
4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) Claim(s) _____ is/are allowed.

6) Claim(s) 1-13 is/are rejected.

7) Claim(s) _____ is/are objected to.

8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) The specification is objected to by the Examiner.

10) The drawing(s) filed on 11 January 2002 is/are: a) accepted or b) objected to by the Examiner.

 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
a) All b) Some * c) None of:
1. Certified copies of the priority documents have been received.
2. Certified copies of the priority documents have been received in Application No. _____.
3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) Notice of References Cited (PTO-892)
2) Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date _____

4) Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____
5) Notice of Informal Patent Application (PTO-152)
6) Other: _____

DETAILED ACTION

1. This Office Action is response to Applicants' response filed on 08/30/2004.
2. Claims 1-13 are pending in this application.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

4. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

5. Claims 1-13 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent No. Patent No. 6,636,862 issued to Lundahl (hereinafter Lundahl) in view of US Patent No. 6,529,891 issued to Heckerman.

With respect to claim 1, Lundahl teaches a) determining a foreground frequency of a bucket within a first cluster (determining the value of a cluster(i) (subscript i) in a cluster model: see fig. 9, col. 20, lines 25-38, col. 6, lines 58-67);

b) determining a background frequency of the bucket with respect to all of the clusters (determining the entire of cluster in the model: col. 8, lines 65-67 and col. 10, lines 1-10);

c) comparing the foreground and background frequencies (comparing the cluster with the entire clusters in the model: col. 12, lines 52-59 and col. 7, lines 52-57); and

d) determining a quality index based on the comparison (the result of the clusters in the model is deriving from the comparison for determining the quality of cluster model: col. 3, lines 40-46 and col. 10, lines 47-67).

Lundahl teaches a method and system for statistical analysis of data model. A data set is containing a number of clusters for determining the appropriate cluster number and analyzing the cluster and entire cluster in the model based the variable of subscript of the cluster. Lundahl does not clearly teach the bucket variable of a cluster variable in the model.

However, Heckerman teaches the variable state that is a value the variable could take (fig. 20, col. 2, lines 15-30; col. 13, lines 3-67 and col. 14, lines 1-28).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to combine the teachings of Lundahl with the teachings of Heckerman, wherein the cluster analysis with the value of variable (see fig. 8), would incorporate the use of variable state for a value of the variable enabling to take, in the same conventional manner as disclosed by Heckerman (col. 13, lines 3-67 and col. 14, lines 1-28). The motivation being for easing to group clustered data into clusters such that belonging to the same cluster have a high degree of similarity.

With respect to claim 2, Lundahl teaches wherein said comparing step further comprises subtracting the relative foreground and background frequencies (col. 4, lines 62-67 and col. 5, lines 1-4 and col. 6, lines 50-58).

With respect to claim 3, Lundahl teaches squaring the result of the comparison (col. 5, lines 9-22).

With respect to claim 4, Lundahl teaches determining an optimal number of clusters and comparing the optimal number of clusters to the actual number of clusters resulting from the clustering date processing operation (see fig. 5, col. 8, lines 65-67, col. 9, lines 1-10, col. 13, lines 55-67 and col. 14, lines 1-2).

With respect to claim 5, Lundahl teaches wherein the optimal number of clusters is determined by a maximum number of buckets for a variable (col. 29, lines 50-67 and col. 30, lines 1-18).

With respect to claim 6, Lundahl teaches wherein the optimal number of clusters is set to a threshold value in case the maximum number of buckets is greater than the threshold value (col. 13, lines 30-55, col. 26, lines 55-67 and col. 27, lines 1-10).

With respect to claims 7-9, Lundahl teaches determining a factor based on the optimal number of clusters and the actual number of clusters (factor analysis: col. 2, lines 10-15 and see fig. 5 and col. 8, lines 65-67 and col. 9, lines 1-10); multiplying the result of the comparison of the relative foreground and background frequencies with the factor (the product of the matrix: col. 4, lines 6-42 and col. 5, lines 24-51); determining a normalizing value being independent of any correlations between fields of the data on which the data processing operation is applied (normalizing the values: col. 12, lines 27-67); normalizing the result of the comparison of the foreground and background frequencies by means of the normalizing value (normalizing the values: col. 12, lines 27-67); comparing the background an imaginary cluster having bucket equal one (comparing the cluster with the entire clusters in the model: col. 12, lines 52-59 and col. 7, lines 52-57); comparing the background frequencies of the buckets with an imaginary cluster having a foreground frequency of the bucket equal to zero (col. 7, lines 52-57); and summing the results of the corresponding comparison values (col. 4, lines 15-28, and the product of the matrix: col. 5, lines 24-51).
With respect to claim 10, Lundahl teaches performing a number of data clustering operation (a given data clustering is performed for estimation an optimal

partitioning of a data set relating to the cluster measure: col. 13, lines 55-67 and col. 14, lines 1-2);

determining a quality index for each result of the data clustering operations (col. 3, lines 40-46); and

selecting the result of the data clustering (col. 19, lines 2-10 and col. 29, lines 8-15).

Lundahl teaches a method and system for statistical analysis of data model. A data set is containing a number of clusters for determining the appropriate cluster number and analyzing the cluster and entire cluster in the model based the variable of subscript of the cluster. Lundahl does not clearly teach the highest quality index as an end result of the data clustering.

However, Heckerman teaches the highest score of the result (the heist score determining the number of states of the cluster variable: col. 20, lines 38-50 and col. 21, lines 45-60).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to combine the teachings of Lundahl with the teachings of Heckerman, wherein the cluster analysis with the value of variable (see fig. 8), would incorporate the use of variable state for a value of the variable enabling to take, in the same conventional manner as disclosed by Heckerman (col. 13, lines 3-67 and col. 14, lines 1-28). The motivation being for easing to group clustered data into clusters such that belonging to the same cluster have a high degree of similarity.

With respect to claim 11, Lundahl teaches a selecting an initial set of clusters (selecting the clustering: col. 14, lines 10-22 and col. 18, lines 45-67); determining a quality index for the clusters (the result of the clusters in the model is deriving from the comparison for determining the quality of cluster model: col. 3, lines 40-46 and col. 10, lines 47-67); and performing a number of iterations to improve the quality index (col. 15, lines 1-26).

Lundahl teaches a method and system for statistical analysis of data model. A data set is containing a number of clusters for determining the appropriate cluster number and analyzing the cluster and entire cluster in the model based the variable of subscript of the cluster. Lundahl does not clearly teach the number of iterations to be performed.

However, Heckerman teaches the number of iterations to be allowed (col. 17, lines 55-67 and col. 18, lines 1-40).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to combine the teachings of Lundahl with the teachings of Heckerman, wherein the cluster analysis with the value of variable (see fig. 8), would incorporate the use of variable state for a value of the variable enabling to take, in the same conventional manner as disclosed by Heckerman (col. 13, lines 3-67 and col. 14, lines 1-28). The motivation being for easing to group clustered data into clusters such that belonging to the same cluster have a high degree of similarity.

However, Lundahl teaches the number of iterations for) and index and the best index value to be selected (col. 13, lines 8-67 and col. 14, lines 1-8).

With respect to claim 12, Lundahl teaches moving at least one record of at least one of the clusters to another cluster (col. 8, lines 65-67 and col. 9, lines 1-20; and col. 13, lines 58-67 and col. 14, lines 1-8); determining the quality index for the modified clusters (the result of the clusters in the model is deriving from the comparison for determining the quality of cluster model: col. 3, lines 40-46 and col. 10, lines 47-67); and using the modified clustered as a new initial set of cluster in case the quality index improved (col. 13, lines 8-67 and col. 14, lines 1-8).

Claim 13 is essentially the same as claim 1 except that it is directed to a computer program product rather than a method (), and is rejected for the same reason as applied to the claim 1 hereinabove.

Contact Information

6. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Anh Ly whose telephone number is (571) 272-4039 or via E-Mail: ANH.LY@USPTO.GOV or fax to (571) 273-4039. The examiner can normally be reached on TUESDAY – THURSDAY from 8:30 AM – 3:30 PM.

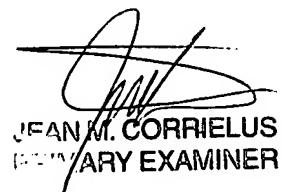
If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Breene, can be reached on (571) 272-4107 or Primary Examiner Jean Corrielus (571) 272-4032.

Any response to this action should be mailed to:

Commissioner of Patents and Trademarks

Washington, D.C. 20231

or faxed to: Central Fax Center (703) 872-9306



JEAN M. CORRIELUS
PRIMARY EXAMINER

ANH LY
JAN. 25th, 2005